SOME DUAL CONDITIONS FOR GLOBAL WEAK SHARP MINIMALITY OF NONCONVEX FUNCTIONS

A. UDERZO

Abstract. Weak sharp minimality is a notion emerged in optimization, whose utility is largeley recognized in the convergence analysis of algorithms for solving extremum problems as well as in the study of the perturbation behaviour of such problems. In the present paper some dual constructions of nonsmooth analysis, mainly related to quasidifferential calculus and its recent developments, are employed in formulating sufficient conditions for global weak sharp minimality. They extend to nonconvex functions a condition, which is known to be valid in the convex case. A feature distinguishing the results here proposed is that they avoid to assume the Asplund property on the underlying space.

1. Introduction and motivations

Given a function $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$, consider the minimization problem

$$(\mathcal{P}) \qquad \min_{x\in X} f(x).$$

Besides the fundamental question on the existence of solutions to $(\mathcal{P})$, i.e. minimizers, an issue of interest is also to single out special behaviours of f able to describe 'how' minima (or, in the global case, the minimum) are attained. Such an issue, which reveals to be crucial for the convergence analysis of algorithms designed to solve problem $(\mathcal{P})$, led to define the notion of weak sharp minima, which is considered here in its global formulation.

In what follows (X, d) denotes a metric space. Given $\alpha \in \mathbb{R} \cup \{+\infty\}$, by $\mathrm{lev}_{\leq\alpha} f = \{x \in X : \ f(x) \leq \alpha\}$ the α sublevel set of f is indicated. The distance of point $x \in X$ from a set $A \subseteq X$ is denoted by $\mathrm{dist}\,(x, A) = \inf_{a\in A} d(x, a)$, with the convention $\mathrm{dist}\,(x, \varnothing) = +\infty$, for every $x \in X$. Set for convenience-sake $\inf_{x\in X} f(x) = \underline{\alpha}$. Throughout the paper it will be assumed that f is bounded from below on X, that is $\underline{\alpha} > -\infty$, unless otherwise stated. Moreover, since the issue under investigation loses its interest if $f \equiv +\infty$, f will be assumed also to be proper, i.e. $\mathrm{dom}\,(f) \neq \varnothing$. $\mathrm{Argmin}(f) = \mathrm{lev}_{\leq\underline{\alpha}} f$ represents the set of all global solutions to problem $(\mathcal{P})$, if any. The acronym l.s.c. (respectively, u.s.c.) abbreviates lower (respectively, upper) semicontinuous, with reference to both functions and set-valued maps.

Definition 1.1. A function $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ defined on a metric space is said to have *global weak sharp minimizers* if $\mathrm{Argmin}(f) \neq \varnothing$ and there exists $\sigma > 0$ such that

$$\sigma \cdot \mathrm{dist}\,(x, \mathrm{Argmin}(f)) \leq f(x) - \underline{\alpha}, \quad \forall x \in X. \tag{1.1}$$

The supremum over all values σ satisfying inequality (1.1) will be called *modulus of global weak sharpness* of f and denoted by $\mathrm{wsha}(f)$.

If, in particular, $\mathrm{Argmin}(f)$ reduces to a singleton $\{\bar{x}\}$, so that inequality (1.1) becomes

$$\sigma d(x, \bar{x}) \leq f(x) - \underline{\alpha}, \quad \forall x \in X, \tag{1.2}$$

such an element $\bar{x}$ is called *sharp minimizer* of f and the related sharpness modulus is denoted by $\mathrm{sha}(f)$.

Date: January 23, 2013.

1991 *Mathematics Subject Classification.* Primary: 49J52; Secondary: 49J53, 90C48.

Key words and phrases. Global weak sharp minimizers, variational principle, strong slope, quasidifferential calculus, lower exhausters, nondegeneracy condition.

For historical remarks on the appearance of weak sharpness in optimization the reader is referred to [3]. As a comment to Definition 1.1 it is worth noting that, in a normed space setting, if f admits a sharp minimizer, then the growth condition (1.2) immediately implies that f is a coercive function, in the sense that

$$\lim_{\|x\|\to\infty} f(x) = +\infty.$$

Nevertheless, the notion of weak sharp minimality is completely independent of coercivity. In fact, a function with global weak sharp minimizers may happen to have sublevel set $\mathrm{Argmin}(f) = \mathrm{lev}_{\leq \underline{\alpha}} f$ not bounded, thereby failing to be coercive. On the other hand, a function such as $x \mapsto \|x\|^2$ is trivially coercive but it admits no global weak sharp minimizers. In other words, the nature of these two properties, although both related to global minimality, appears to be totally different.

Another aspect of Definition 1.1 to be remarked is that global weak sharp minimality is a property of function f itself, not a property qualifying some of its global minimizers. This makes global weak sharp minimality different from its local counterpart. Since, whenever it exists, the global minimum of a function is unique, in the global case it makes no sense to speak of 'weak sharp minim*a*'. For this reason, in formulating Definition 1.1, the term 'weak sharp minimizers' has been preferred to the one largely employed in the existent literature, namely 'weak sharp minima'.

Further features of weak sharp minimizers come up when X has some more structure. Let us mention in this concern that, if X is a Banach space and f is a convex function not constant on X, then the occurence of global weak sharp minimizers is incompatible with the differentiability of f (see, for instance, Proposition 2.4 in [24]).

In what follows several exemplary contexts from optimization and variational analysis are illustrated, in which the concept of weak sharp minimizer naturally occurs. The resulting insight should enlighten the crucial role played by the notion under consideration in the mentioned disciplines.

Example 1.2. A function $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ defined on a metric space is said to have a *global error bound* if there exists $\zeta > 0$ such that

$$\mathrm{dist}\,(x, \mathrm{lev}_{\leq \alpha} f) \leq \zeta [f(x) - \alpha]_+, \quad \forall \alpha \in \mathbb{R},\ \forall x \in\ X,$$

where, given any real r, it is $[r]_+ = \max\{r, 0\}$. Thus, the property of having global weak sharp minimizers amounts to the existence of a global error bound in the particular case in which $\alpha = \underline{\alpha}$, with $\zeta \geq \mathrm{wsha}(f)^{-1}$. Notice that, with the convention here adopted $\mathrm{dist}\,(x, \varnothing) = +\infty$, the existence of a global error bound for a function bounded from below entails that $\mathrm{Argmin}(f) \neq \varnothing$.

Example 1.3. Let $F : X \longrightarrow 2^X$ a set-valued map defined and taking closed and bounded values on the same metric space (X, d). Suppose that F is a contraction of X, i.e. there exists $\kappa \in [0, 1)$ such that

$$\mathrm{Haus}(F(x_1), F(x_2)) \leq \kappa d(x_1, x_2), \quad \forall x_1,\, x_2 \in X,$$

where $\mathrm{Haus}(A, B)$ indicates the Hausdorff distance of the closed and bounded sets A and B. If X is metrically complete the displacement function associated to F, namely function $\phi_F : X \longrightarrow [0, +\infty)$ defined as

$$\phi_F(x) = \mathrm{dist}\,(x, F(x))\,,$$

turns out to admit global weak sharp minimizers. Actually, they coincide with the fixed points of F. Indeed, if denoting by $\mathrm{Fix}(F)$ the set of all fixed points of F, it is known from the Covitz-Nadler theorem on fixed points that $\mathrm{Fix}(F) \neq \varnothing$ and, for every $\kappa' \in (\kappa, 1)$, it results in

$$\mathrm{dist}\,(x, \mathrm{Fix}(F)) \leq \frac{\mathrm{dist}\,(x, F(x))}{1 - \kappa'}, \quad \forall x \in X, \tag{1.3}$$

(see [9, 23]). Thus, one deduces that $\mathrm{wsha}(\phi_F) \geq 1 - \kappa$. Notice that, generally speaking, function ϕ_F fails to admit a global sharp minimizer. Nevertheless, since $\mathrm{Fix}(F)$ reduces to a singleton whenever F is single-valued on X according to the Banach-Caccioppoli contraction principle, then in the latter case ϕ_F admits a global sharp minimizer and $\mathrm{sha}(\phi_F) \geq 1 - \kappa$.

Example 1.4. According to [21] a set-valued map $F : Y \longrightarrow 2^X$ between metric spaces is said to be *calm* at $(y_0, x_0) \in \mathrm{Gr}\, F$ if there exist $r > 0$ and $\gamma > 0$ such that

$$\mathrm{dist}\,(x, F(y_0)) \leq \gamma d(y, y_0), \quad \forall x \in F(y) \cap \mathrm{B}\,(x_0, r)\,,\ \forall y \in \mathrm{B}\,(y_0, r)\,,$$

where $\mathrm{B}(x,r)=\mathrm{lev}_{\le r} d(\cdot,x)$. Let $f : X \longrightarrow \mathbb{R}\cup\{+\infty\}$ be a given function. Take $Y=\mathbb{R}$, $y_0=\underline{\alpha}$ and consider the set-valued map $\mathrm{Lev}\, f:\mathbb{R}\longrightarrow 2^X$ associated to f as follows

$$\mathrm{Lev}\, f(\alpha)=\mathrm{lev}_{\le\alpha} f.$$

Assume that $\mathrm{Argmin}(f)\neq\varnothing$ and that $\mathrm{Lev}\, f$ is calm at each pair $(\underline{\alpha},x)$, where $x\in\mathrm{Argmin}(f)$, and with $r=+\infty$ and a uniform constant γ. Then, being

$$\mathrm{dist}\,(x,\mathrm{Argmin}(f))=\mathrm{dist}\,(x,\mathrm{Lev}\, f(\underline{\alpha}))\le\gamma|f(x)-\underline{\alpha}|=\gamma(f(x)-\underline{\alpha}),\quad \forall x\in\mathrm{Lev}\, f(f(x))=X,$$

one sees that under the above global calmness condition function f has global weak sharp minimizers.

An useful achievement of modern variational analysis is that the calmness property of a given map captures a Lipschitz behaviour that generalizes the metric regularity of its inverse, known as metric subregularity. This fact leads to the next context in which weak sharp minimality arises.

Example 1.5. Let $f : X \longrightarrow \mathbb{R}\cup\{+\infty\}$ be a given function defined on a metric space. According to [19], f is said to be *metrically subregular* at $x_0\in X$ for $\alpha_0\in\mathbb{R}$ if there exist $r>0$ and $\zeta>0$ such that

$$\mathrm{dist}\left(x,f^{-1}(\alpha_0)\right)\le\zeta|f(x)-\alpha_0|,\quad \forall x\in\mathrm{B}\,(x_0,r)\,.$$

From the above inequality, one readily sees that any function f, which is metrically subregular at each point of X for $\underline{\alpha}$, with the same constant ζ, admits global weak sharp minimizers.

Example 1.6. A given function $f : X \longrightarrow \mathbb{R}\cup\{+\infty\}$ defined on a topological space X is said to be *Tikhonov well-posed* if $\mathrm{Argmin}(f)=\{\bar{x}\}$ and for every sequence $(x_n)_{n\in\mathbb{N}}$ in X, such that $f(x_n)\to\underline{\alpha}$ as $n\to\infty$ (minimizing sequence), one has $x_n\to\bar{x}$. It is clear that, in a metric space setting, whenever f admits a global sharp minimizer it is also Tikhonov well-posed. The converse is evidently false, in general. If f has global weak sharp minimizers without a global sharp minimizer, from inequality (1.1) one easily deduces that any minimizing sequence is metrically attracted by $\mathrm{Argmin}(f)$. Thus, global weak sharp minimality leads to a set valued-like generalization of the Tikhonov well-posedness.

In consideration of its theoretical and computational relevance, weak sharp minimality has been the subject of manifold investigations within optimization and variational analysis (see, for instance, [2, 3, 4, 5, 6, 24, 28, 31]). A topic considered in several of them is the problem of finding methods to detect the occurence of weak sharp minimality. The present paper intends to focus on this topic. Following a recognized line of reasearch, the goal of the analysis here proposed is to extend a known condition valid for convex functions beyond the realm of convexity. Since, as already mentioned, weak sharp minimality is incompatible with differentiability, this task is pursued by making use of nonsmooth analysis tools. In particular, sufficient conditions for global weak sharp minimality are presented in the case of problems $(\mathcal{P})$ with objective function quasidifferentiable or with generalized derivatives admitting lower exhausters. In both cases, the key role is played by respective nondegeneracy conditions involving subdifferential-like dual constructions. A feature of the present analysis is that the resulting conditions are achieved as a consequence of a general analysis conducted in a metric space setting.

The material exposed in rest of the paper is arranged as follows. In Section 2 a characterization and a sufficient condition for global weak sharp minimality are established and discussed in a metric space setting. In Section 3 some material from nonsmooth analysis, which is needed in order to formulate nondegeneracy conditions, is briefly recalled and some related ancillary result is proved. Section 4 is reserved to present and comment the main results of the paper. They consider both unconstrained and variously constrained extremum problems.

2. Global weak sharp minimality in metric spaces

Even though metric space is in structure too poor for certain applications, nevertheless it should be the proper environment where to analyze the notion of weak sharp minimality 'juxta propria principia', in consideration of the purely metric nature of its definition. In fact, all dual conditions presented in Section 4 will be derived from basic results established in the present section. Let us start with a characterization for f to have global weak sharp minimizers, which relies on the behaviour of sublevel sets of f.

Theorem 2.1. *Let (X,d) be a complete metric space. A function $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ l.s.c. on X admits global weak sharp minimizers with modulus* $\mathrm{wsha}(f) \geq \tau > 0$ *iff*

$$\tau \cdot \sup_{\alpha \in (\underline{\alpha}, +\infty)} \mathrm{dist}\,(x, \mathrm{lev}_{\leq \alpha} f) \leq f(x) - \underline{\alpha}, \quad \forall x \in X. \tag{2.1}$$

Proof. Suppose first that f does admit global weak sharp minimizers, having modulus $\mathrm{wsha}(f)$. Then, since it is

$$\varnothing \neq \mathrm{Argmin}(f) \subseteq \mathrm{lev}_{\leq \alpha} f, \quad \forall \alpha \in (\underline{\alpha}, +\infty),$$

consequently, for every $\sigma \in (0, \mathrm{wsha}(f))$ and $\alpha \in (\underline{\alpha}, +\infty)$, it holds

$$\sigma \cdot \mathrm{dist}\,(x, \mathrm{lev}_{\leq \alpha} f) \leq \sigma \cdot \mathrm{dist}\,(x, \mathrm{Argmin}(f)) \leq f(x) - \underline{\alpha}, \quad \forall x \in X,$$

whence inequality (2.1) follows at once.

Conversely, suppose condition (2.1) to hold true. Fix an arbitrary $x_0 \in X \backslash \mathrm{Argmin}(f)$. Without loss of generality it is possible to assume that $x_0 \in \mathrm{dom}\,(f)$. Indeed, in the case $x_0 \notin \mathrm{dom}\,(f)$ inequality (1.1) would be automatically true (remember that $\mathrm{dom}\,(f) \neq \varnothing$). Take an arbitrary $\sigma \in (0, \tau)$. Since it is $f(x_0) \leq \underline{\alpha} + (f(x_0) - \underline{\alpha})$, then by virtue of the Ekeland variational principle, corresponding to the value

$$\lambda = \frac{f(x_0) - \underline{\alpha}}{\sigma},$$

there exists $x_\lambda \in X$ with the properties:

$$f(x_\lambda) \leq f(x_0), \tag{2.2}$$

$$d(x_\lambda, x_0) \leq \lambda, \tag{2.3}$$

$$f(x_\lambda) < f(x) + \sigma d(x, x_\lambda), \quad \forall x \in X \backslash \{x_\lambda\}. \tag{2.4}$$

Inequality (2.2) says that $x_\lambda \in \mathrm{dom}\,(f)$. Let us show that such an x_λ is a global weak sharp minimizer of f. Indeed, assume ab absurdo that $x_\lambda \notin \mathrm{Argmin}(f)$, so $\underline{\alpha} < f(x_\lambda)$. Then, condition (2.1) being valid by hypothesis, from the fact that

$$\mathrm{dist}\,(x_\lambda, \mathrm{lev}_{\leq \alpha} f) > 0, \quad \forall \alpha \in (\underline{\alpha}, f(x_\lambda)),$$

one obtains

$$\sigma \cdot \sup_{\alpha \in (\underline{\alpha}, f(x_\lambda))} \mathrm{dist}\,(x_\lambda, \mathrm{lev}_{\leq \alpha} f) < \tau \cdot \sup_{\alpha \in (\underline{\alpha}, +\infty)} \mathrm{dist}\,(x_\lambda, \mathrm{lev}_{\leq \alpha} f) \leq f(x_\lambda) - \underline{\alpha}.$$

This means that it is possible to find $\epsilon > 0$ such that

$$\sigma \cdot \sup_{\alpha \in (\underline{\alpha}, f(x_\lambda))} \mathrm{dist}\,(x_\lambda, \mathrm{lev}_{\leq \alpha} f) < f(x_\lambda) - \underline{\alpha} - \epsilon.$$

In particular, as it is $\underline{\alpha} < \underline{\alpha} + \epsilon < f(x_\lambda)$, it holds

$$\sigma \cdot \mathrm{dist}\left(x_\lambda, \mathrm{lev}_{\leq \underline{\alpha} + \epsilon} f\right) < f(x_\lambda) - \underline{\alpha} - \epsilon.$$

From the last inequality it follows that there exists $x_\epsilon \in \mathrm{lev}_{\leq \underline{\alpha} + \epsilon} f$ such that

$$\sigma d(x_\lambda, x_\epsilon) < f(x_\lambda) - \underline{\alpha} - \epsilon.$$

Notice that $x_\epsilon \in X \backslash \{x_\lambda\}$, because x_λ does not belong to $\mathrm{lev}_{\leq \underline{\alpha} + \epsilon} f$. Thus, by recalling inequality (2.4), one finds

$$f(x_\lambda) < f(x_\epsilon) + \sigma d(x_\epsilon, x_\lambda) < f(x_\epsilon) + f(x_\lambda) - \underline{\alpha} - \epsilon \leq f(x_\lambda),$$

which clearly leads to an absurdum. This argument hence shows that $x_\lambda \in \mathrm{Argmin}(f) \neq \varnothing$. Again, by virtue of inequality (2.3), by recalling the chosen value of λ one obtains

$$\mathrm{dist}\,(x_0, \mathrm{Argmin}(f)) \leq d(x_0, x_\lambda) \leq \frac{f(x_0) - \underline{\alpha}}{\sigma},$$

whence

$$\sigma \cdot \mathrm{dist}\,(x_0, \mathrm{Argmin}(f)) \leq f(x_0) - \underline{\alpha}. \tag{2.5}$$

Since the validity of inequality (2.5) has been proved for every $\sigma \in (0, \tau)$, one can deduce that

$$\tau \cdot \mathrm{dist}\,(x_0, \mathrm{Argmin}(f)) \leq f(x_0) - \underline{\alpha}.$$

By arbitrariness of x_0, all requirements of Definition 1.1 appear now to be fulfilled. Therefore the proof is complete. □

Remark 2.2. It is helpful to mention that, since for every $x \in X$ function $\alpha \mapsto \mathrm{dist}\,(x, \mathrm{lev}_{\leq \alpha} f)$ is monotone decreasing on $[\underline{\alpha}, +\infty)$, then condition (2.1) can be equivalently rewritten as

$$\tau \cdot \lim_{\alpha \to \underline{\alpha}^+} \mathrm{dist}\,(x, \mathrm{lev}_{\leq \alpha} f) \leq f(x) - \underline{\alpha}, \quad \forall x \in X.$$

In fact, Theorem 2.1 has been devised as a modification of a similar result presented in [24], where metric completeness has been exploited through the shrinking ball property, instead of through the Ekeland's principle.

A basic sufficient condition for global weak sharp minimality can be formulated in terms of strong slope. This is a variational analysis tool originally proposed in [10] for quite different purposes, whose utility has been demonstrated in several circumstances (see, for instance, [1, 20]). Given a function $f : X \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ defined on a metric space, by *strong slope* of f at $\bar{x} \in \mathrm{dom}\,(f)$ the real-extended value

$$|\nabla f|(\bar{x}) = \begin{cases} 0, & \text{if } \bar{x} \text{ is a local minimizer for } f, \\ \displaystyle\limsup_{x \to \bar{x}} \frac{f(\bar{x}) - f(x)}{d(x, \bar{x})}, & \text{otherwise,} \end{cases}$$

is meant. In the case $\bar{x} \notin \mathrm{dom}\,(f)$, set $|\nabla f|(\bar{x}) = +\infty$.

Example 2.3. The above recalled tool takes a form more appealing from the computational viewpoint when X and f possess more structure. For instance, if X is a normed vector space, then for any function f Fréchet differentiable at $x \in X$, with Fréchet derivative $\hat{\mathrm{D}}f(x)$, it holds

$$|\nabla f|(x) = \|\hat{\mathrm{D}}f(x)\|,$$

(what motivates the notation in use). If, in the same space, f is convex and subdifferentiable at x, it holds

$$|\nabla f|(x) = \mathrm{dist}\,(\mathbf{0}^*, \partial f(x)),$$

where $\partial f(x)$ denotes the subdifferential of f at x in the sense of convex analysis and $\mathbf{0}^*$ stands for the null vector of the dual space (see Section 3).

Theorem 2.4. *Let (X, d) be a complete metric space and let $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ be l.s.c. on X. If there exists $\tau > 0$ such that*

$$|\nabla f|(X \backslash \mathrm{Argmin}(f)) \subseteq [\tau, +\infty), \tag{2.6}$$

then f admits global weak sharp minimizers and $\mathrm{wsha}(f) \geq \tau$.

Proof. Observe first that if $\mathrm{Argmin}(f) = X$ the thesis becomes trivial. So, fix arbitrarily x_0 and σ in $(X \backslash \mathrm{Argmin}(f)) \cap \mathrm{dom}\,(f)$ and $(0, \tau)$, respectively. In the case $x_0 \notin \mathrm{dom}\,(f)$ there is nothing to be proved. According to the Ekeland variational principle, corresponding to $\lambda = (f(x_0) - \underline{\alpha})/\sigma$ there exists $x_\lambda \in \mathrm{dom}\,(f)$ such that

$$d(x_\lambda, x_0) \leq \lambda \tag{2.7}$$

and

$$f(x_\lambda) < f(x) + \sigma d(x, x_\lambda), \quad \forall x \in X \backslash \{x_\lambda\}.$$

The last inequality implies that for every $\delta > 0$ it results in

$$\sup_{x \in \mathrm{B}(x_\lambda, \delta) \backslash \{x_\lambda\}} \frac{f(x_\lambda) - f(x)}{d(x, x_\lambda)} \leq \sup_{x \in X \backslash \{x_\lambda\}} \frac{f(x_\lambda) - f(x)}{d(x, x_\lambda)} \leq \sigma,$$

wherefrom it follows

$$|\nabla f|(x_\lambda) \leq \sigma. \tag{2.8}$$

Now observe that it must be $x_\lambda \in \mathrm{Argmin}(f)$. Otherwise, if it were $x_\lambda \in X \backslash \mathrm{Argmin}(f)$, inequality (2.8) would contradict condition (2.6). From inequality (2.7) one obtains

$$\sigma \mathrm{dist}\,(x_0, \mathrm{Argmin}(f)) \leq f(x_0) - \underline{\alpha}.$$

The arbitrariness of x_0 and σ in their respective domains allows one to complete the proof. □

The next example shows that, in contrast to the previous result, Theorem 2.4 is far removed from being a characterization of global weak sharp minimality.

Example 2.5. Let $X = \mathbb{R}$ be equipped with its usual (Euclidean) metric structure. Consider function $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined by

$$f(x) = \begin{cases} 0, & \text{if } x \in (-\infty, 0], \\ x + \dfrac{1}{x+1}, & \text{otherwise.} \end{cases}$$

Clearly, f is l.s.c. and bounded from below on $\mathbb{R}$. Here $\underline{\alpha} = 0$ and $\mathrm{Argmin}(f) = (-\infty, 0]$. Thus, since it results in

$$\mathrm{dist}\,(x, \mathrm{Argmin}(f)) = [x]_+ \le x + \frac{1}{x+1}, \quad \forall x \in (0, +\infty),$$

one deduces that f admits global weak sharp minimizers, with $\mathrm{wsha}(f) \ge 1$. Nonetheless, being f differentiable on $\mathbb{R}\backslash\{0\}$, by taking into account Example 2.3, one finds

$$|\nabla f|(x) = 1 - \frac{1}{(x+1)^2}, \quad \forall x \in (0, +\infty),$$

what makes condition (2.6) evidently violated.

Remark 2.6. As, according to the definition of strong slope, it is $|\nabla f|(\mathrm{Argmin}(f)) = \{0\}$, condition (2.6) entails that, if f is not constant, function $|\nabla f| : X \longrightarrow [0, +\infty]$ can not be u.s.c. (and hence continuous) at each point $\bar{x} \in \mathrm{Argmin}(f)$, which is an accumulation point of $X\backslash\mathrm{Argmin}(f)$. Indeed, for any sequence $(x_n)_{n\in\mathbb{N}}$, with $x_n \in X\backslash\mathrm{Argmin}(f)$ and $x_n \to \bar{x}$ as $n \to \infty$, one finds

$$\limsup_{n\to\infty} |\nabla f|(x_n) \ge \liminf_{n\to\infty} |\nabla f|(x_n) \ge \tau > 0 = |\nabla f|(\bar{x}).$$

Theorem 2.4 enables one to easily derive a global error bound for the solution set of an inequality/equality system. Given $g : X \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ and $h : X \longrightarrow \mathbb{R} \cup \{\pm\infty\}$, for any α, $\beta \in \mathbb{R}$ define

$$\Omega_{\alpha,\beta} = \mathrm{lev}_{\le\alpha}\, g \cap h^{-1}(\beta).$$

Corollary 2.7. *Let (X, d) be a complete metric space, let $g : X \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ be l.s.c. on X, let $h : X \longrightarrow \mathbb{R}$ be continuous on X, both not necessarily bounded from below, and let α, $\beta \in \mathbb{R}$. If $\Omega_{\alpha,\beta} \neq \varnothing$ and there exists $\tau > 0$ such that*

$$|\nabla([g - \alpha]_+ + |h - \beta|)|(X\backslash\Omega_{\alpha,\beta}) \subseteq [\tau, +\infty),$$

then the following error bound holds

$$\mathrm{dist}\,(x, \Omega_{\alpha,\beta}) \le \tau^{-1}\left([g(x) - \alpha]_+ + |h(x) - \beta|\right), \quad \forall x \in X.$$

Proof. Set $f = [g - \alpha]_+ + |h - \beta|$. Since by hypothesis $\Omega_{\alpha,\beta} \neq \varnothing$, then it is $\inf\limits_{x\in X} f(x) = 0$. Such an infimum is actually attained and one has $\Omega_{\alpha,\beta} = \mathrm{Argmin}(f)$. By definition, under the assumptions made, function f is l.s.c. and obviously bounded from below. Thus, one is in a position to apply Theorem 2.4. □

3. Selected elements of Nonsmooth Analysis

Throughtout the current section, $(\mathbb{X}, \|\cdot\|)$ denotes a real Banach space. The (topological) dual space of $\mathbb{X}$ is marked by $\mathbb{X}^*$, with $\mathbb{X}^*$ and $\mathbb{X}$ being paired in duality by the bilinear form $\langle \cdot, \cdot \rangle : \mathbb{X}^* \times \mathbb{X} \longrightarrow \mathbb{R}$. The null vector of $\mathbb{X}$ is indicated by $\mathbf{0}$, while the null functional by $\mathbf{0}^*$. $\mathbb{B}^*$ denotes the unit ball of $\mathbb{X}^*$, while $\mathbb{S}$ denotes the unit sphere in $\mathbb{X}$. The convention $A + \varnothing = \varnothing$ is adopted for any subset A of a vector space. The commutative semigroup of all (Lipschitz) continuous sublinear functions defined on $\mathbb{X}$ is denoted by $\mathcal{S}(\mathbb{X})$. A starting point for entering the subsequent dual constructions is the semigroup isomorphism between $\mathcal{S}(\mathbb{X})$ and the semigroup of all nonempty convex and weak* compact subsets of $\mathbb{X}^*$, denoted by $\mathcal{K}(\mathbb{X}^*)$. Such isomorphism is known as *Minkowski-Hörmander duality* and is represented by the subdifferential map $\partial : \mathcal{S}(\mathbb{X}) \longrightarrow \mathcal{K}(\mathbb{X}^*)$ as follows

$$\partial(h) = \partial h(\mathbf{0}), \quad h \in \mathcal{S}(\mathbb{X}).$$

Clearly, $\partial^{-1} : \mathcal{K}(\mathbb{X}^*) \longrightarrow \mathcal{S}(\mathbb{X})$ can be defined via the support function to a given subset, namely

$$\partial^{-1}(A) = \varsigma(\cdot, A), \quad A \in \mathcal{K}(\mathbb{X}^*),$$

where $\varsigma(x, A) = \max\limits_{x^*\in A} \langle x^*, x \rangle$.

3.1. **Quasidifferentiable functions and Demyanov difference.** Given a function $f:\mathbb{X}\longrightarrow\mathbb{R}\cup\{\pm\infty\}$ and $\bar{x}\in\mathrm{dom}\,(f)$, let $f'(\bar{x};v)$ denote the directional derivative of f at $\bar{x}$ in the direction $v\in\mathbb{X}$. According to [14], function f is said to be *quasidifferentiable* (for short, *q.d.*) at $\bar{x}$ if it admits directional derivative at $\bar{x}$ in all directions and there exist two elements $\underline{f},\,\overline{f}\in\mathcal{S}(\mathbb{X})$ such that

$$f'(\bar{x};v)=\underline{f}(v)-\overline{f}(v),\quad \forall v\in\mathbb{X}.$$

In the light of the Minkowski-Hörmander duality this amounts to say that the following dual representation is valid

$$f'(\bar{x};v)=\varsigma(v,\partial\underline{f}(\mathbf{0}))-\varsigma(v,\partial\overline{f}(\mathbf{0})),\quad \forall v\in\mathbb{X}. \tag{3.1}$$

Clearly, the representation in (3.1) of $f'(\bar{x};\cdot)$, as well as the previous one in terms of $\mathcal{S}(\mathbb{X})$, is by no means unique. This is not a serious drawback, because every pair of elements of $\mathcal{K}(\mathbb{X}^*)$ representing $f'(\bar{x};\cdot)$ belongs to the same class with respect to an equivalence relation $\sim$ defined on $\mathcal{K}(\mathbb{X}^*)\times\mathcal{K}(\mathbb{X}^*)$, according to which $(A,B)\sim(C,D)$ if $A+D=B+C$. The $\sim$-equivalence class containing the pair $(\partial\underline{f}(\mathbf{0}),-\partial\overline{f}(\mathbf{0}))$ is called *quasidifferential* of f at $\bar{x}$ and will be denoted in further constructions by $\mathcal{D}f(\bar{x})$. Any pair in the class $\mathcal{D}f(\bar{x})$ will be henceforth indicated by $(\underline{\partial}f(\bar{x}),-\overline{\partial}f(\bar{x}))$, so

$$f'(\bar{x};v)=\varsigma(v,\underline{\partial}f(\bar{x}))-\varsigma(v,-\overline{\partial}f(\bar{x})),\quad \forall v\in\mathbb{X},\qquad\text{and}\qquad \mathcal{D}f(\bar{x})=[\underline{\partial}f(\bar{x}),-\overline{\partial}f(\bar{x})]_\sim.$$

In the early 80-ies a complete calculus for quasidifferentiable functions has been developed, which finds a geometric counterpart in the calculus for $\sim$-equivalence classes of pairs in $\mathcal{K}(\mathbb{X}^*)\times\mathcal{K}(\mathbb{X}^*)$ (see [15, 16, 25]). This, along with a notable computational tractability of the resulting constructions, made such approach a recognized and successful subject within nonsmooth analysis.

The next step towards the setting of analysis tools in use in the subsequent section requires the introduction of a difference operation in $\mathcal{K}(\mathbb{X}^*)$. As illustrated in several works (see, for instance, [16, 25, 26, 27]), such a task can be accomplished following different approaches. For the purposes of the present investigations, the following notion seems to be adequate.

Definition 3.1. The inner operation $\,\underline{\triangle}\,:\mathcal{K}(\mathbb{X}^*)\times\mathcal{K}(\mathbb{X}^*)\longrightarrow\mathcal{K}(\mathbb{X}^*)$, defined by

$$A\,\underline{\triangle}\,B=\partial^\circ(\varsigma(\cdot,A)-\varsigma(\cdot,B))(\mathbf{0}), \tag{3.2}$$

where ∂° denotes the Clarke subdifferential operator (see [8, 16]), is called *Demyanov difference* of A and B.

Remark 3.2. Notice that, since function $\varsigma(\cdot,A)-\varsigma(\cdot,B)$ is Lipschitz continuous, the expression in (3.2) actually makes sense. In other words, $A\,\underline{\triangle}\,B$ never gives $\varnothing$. It is to be mentioned that the definition proposed above is not the original one, as it was introduced in [11]. The latter was formulated in the more particular setting of finite dimensional Euclidean spaces. It is nonetheless relevant to recall it in detail in as much as this should offer insights into the potential of such tool. Fixed a point $x\in\mathbb{R}^n$ and an element $A\in\mathcal{K}(\mathbb{R}^n)$, the set

$$\Phi_x(A)=\{a\in A:\langle a,x\rangle=\varsigma(x,A)\}$$

is called *max-face* of A generated by x. The dual representation

$$\varsigma(x,A)=\max_{a\in\partial\varsigma(\cdot,A)(x)}\langle a,x\rangle$$

shows that $\Phi_x(A)=\partial\varsigma(\cdot,A)(x)$. Now, from the differentiability theory of convex functions it is well known that $\partial\varsigma(\cdot,A)(x)$ is a singleton iff function $\varsigma(\cdot,A)$ admits gradient $\nabla\varsigma(\cdot,A)(x)$ at x. Observe that in such a circumstance the max-face $\Phi_x(A)$ reduces to what is called an exposed point of A generated by the hyperplane $\langle\cdot\,,x\rangle$. Then, according to the Rademacher's theorem, setting

$$M_A=\{x\in\mathbb{R}^n:\Phi_x(A)\text{ is a singleton}\}$$

and denoting by μ the Lebesgue measure on $\mathbb{R}^n$, it results in

$$\mu(\mathbb{R}^n\backslash M_A)=0,$$

i.e. M_A is a Lebesgue full-measure set. Now recall that, given a locally Lipschitz function $\varphi:\mathbb{R}^n\longrightarrow\mathbb{R}$, a point $\bar{x}\in\mathbb{R}^n$ and a full-measure set M, at the points of which φ admits gradient, it holds

$$\partial^\circ\varphi(\bar{x})=\mathrm{conv}\,\{v\in\mathbb{R}^n:\ \exists(x_n)_{n\in\mathbb{N}},\ x_n\in M,\ x_n\to\bar{x},\ v=\lim_{n\to\infty}\nabla\varphi(x_n)\},$$

where conv denotes the convex hull of a given set. Consequently, letting $M_{A,B}$ be a full-measure subset of $\mathbb{R}^n$ where both functions $\varsigma(\cdot, A)$ and $\varsigma(\cdot, B)$ admits gradient, it is possible to define

$$A \,\underline{\triangle}\, B = \mathrm{cl\,conv}\,\{\nabla\varsigma(\cdot, A)(x) - \nabla\varsigma(\cdot, B)(x) :\ x \in M_{A,B}\},$$

where cl conv indicates convex closure. This finite dimensional reading of $\underline{\triangle}$ allows one for the following constructive view of $A \,\underline{\triangle}\, B$: such set turns out to consist of all differences of points, respectively in A and B, which are exposed by the same hyperplane $\langle \cdot, x\rangle$, with x varying in $M_{A,B}$.

The below lemma collects those properties of the Demyanov difference that will be exploited in the sequel. Their proofs, as well as additional material and further discussion on this topic, can be found in [16, 25, 26, 27].

Lemma 3.3. *(1) For every A, B, C, $D \in \mathcal{K}(\mathbb{X}^*)$, if $(A, B) \sim (C, D)$ then $A \,\underline{\triangle}\, B = C \,\underline{\triangle}\, D$;*
(2) for every $A \in \mathcal{K}(\mathbb{X}^*)$, it holds $A \,\underline{\triangle}\, A = \{\mathbf{0}^*\}$;
(3) for every A, $B \in \mathcal{K}(\mathbb{X}^*)$, if $B \subseteq A$, then $\mathbf{0}^* \in A \,\underline{\triangle}\, B$;
(4) for every A, B, C, $D \in \mathcal{K}(\mathbb{X}^*)$, the inclusion $(A + B) \,\underline{\triangle}\, (C + D) \subseteq (A \,\underline{\triangle}\, C) + (B \,\underline{\triangle}\, D)$ holds true;
(5) for every $A \in \mathcal{K}(\mathbb{X}^*)$, it holds $A \,\underline{\triangle}\, \{\mathbf{0}^*\} = A$;
(6) for every A, $B \in \mathcal{K}(\mathbb{X}^*)$, it is $A \,\underline{\triangle}\, B \subseteq A - B$.

Given a function $f : \mathbb{X} \longrightarrow \mathbb{R}$, suppose that f is q.d. at each point of $\mathbb{X}$, with $\mathcal{D}f(x) = [\underline{\partial} f(x), -\overline{\partial} f(x)]_\sim$. By combining the quasidifferential pairs of f with Demyanov difference the following generalized derivative construction $\mathcal{D}^{\,\triangleq} f : \mathbb{X} \longrightarrow \mathcal{K}(\mathbb{X}^*)$, which will be employed to establish a weak sharp minimality condition, is obtained:

$$\mathcal{D}^{\,\triangleq} f(x) = \underline{\partial} f(x) \,\underline{\triangle}\, (-\overline{\partial} f(x)).$$

The use of the above construction is not new in nonsmooth analsysis (as an example of different employments, see for instance [18, 30]).

Remark 3.4. (1) Notice that, by virtue of Lemma 3.3(1), map $\mathcal{D}^{\,\triangleq} f$ does not depend on particular representations of $\mathcal{D}f$. Since, as one immediately checks, one finds $\mathcal{D}^{\,\triangleq} h(\mathbf{0}) = \partial(h)$ whenever $h \in \mathcal{S}(\mathbb{X})$, construction $\mathcal{D}^{\,\triangleq}$ can be regarded as an extension of the Minkowski-Hörmander duality.

(2) In force of Lemma 3.3(3) it is readily seen that, whenever $\bar{x} \in \mathbb{X}$ is a local minimizer of f, it must be $\mathbf{0}^* \in \mathcal{D}^{\,\triangleq} f(\bar{x})$.

(3) By employing Lemma 3.3(4) and the well-known sume rule for quasidifferentials (see, for instance, [16]), it is possible to prove that, given two functions $f : \mathbb{X} \longrightarrow \mathbb{R}$ and $g : \mathbb{X} \longrightarrow \mathbb{R}$, both q.d. at $x \in \mathbb{X}$, the following inclusion holds

$$\mathcal{D}^{\,\triangleq}(f + g)(x) \subseteq \mathcal{D}^{\,\triangleq} f(x) + \mathcal{D}^{\,\triangleq} g(x).$$

(4) By employing the calculus rule for quasidifferentials of functions by scalars, it is possible to prove that, given a function $f : \mathbb{X} \longrightarrow \mathbb{R}$ q.d. at $x \in \mathbb{X}$ and a scalar $\alpha \geq 0$, the following equality holds

$$\mathcal{D}^{\,\triangleq}[\alpha f](x) \subseteq \alpha \mathcal{D}^{\,\triangleq} f(x).$$

(5) From Lemma 3.3(6) it is evident that, if a function $f : \mathbb{X} \longrightarrow \mathbb{R}$ is q.d. at x, with $\mathcal{D}f(x) = [\underline{\partial} f(x), -\overline{\partial} f(x)]_\sim$, then it is always $\mathcal{D}^{\,\triangleq} f(x) \subseteq \underline{\partial} f(x) + \overline{\partial} f(x)$. Immediate examples show that this inclusion can be strict. Notice that the right side term of it depends on the pair in $\mathcal{K}(\mathbb{X}^*) \times \mathcal{K}(\mathbb{X}^*)$ representing $\mathcal{D}f(x)$.

Example 3.5. Let $f : \mathbb{X} \longrightarrow \mathbb{R}$ be a continuous convex function. Since in this case f is q.d. at each point $x \in \mathbb{X}$ with $\mathcal{D}f(x) = [\partial f(x), \{\mathbf{0}^*\}]_\sim$, for such kind of function owing to Lemma 3.3(5) one obtains

$$\mathcal{D}^{\,\triangleq} f(x) = \partial f(x) \,\underline{\triangle}\, \{\mathbf{0}^*\} = \partial f(x).$$

When, in particular, f happens to be Gâteaux differentiable at x, with Gâteaux derivative $\nabla f(x)$, one has $\mathcal{D}^{\,\triangleq} f(x) = \{\nabla f(x)\}$. Therefore, if $\mathbb{X}$ is a separable Banach space, map $\mathcal{D}^{\,\triangleq} f$ is single valued on a G_δ dense subset of $\mathbb{X}$, on account of Mazur theorem.

In view of the employment of construction $\mathcal{D}^{\,\triangleq}$ in the analysis of global weak sharp minimality through the condition established in Theorem 2.4, the next estimate, already obtained in [30] within a more general argument, is needed.

Lemma 3.6. *Given a function $f : \mathbb{X} \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ defined on a normed vector space $(\mathbb{X}, \|\cdot\|)$, suppose that f is q.d. at $\bar{x} \in \mathrm{dom}\,(f)$. Then, the following inequality holds*

$$|\nabla f|(\bar{x}) \geq \mathrm{dist}\left(\mathbf{0}^*, \mathcal{D}^{\triangle} f(\bar{x})\right). \tag{3.3}$$

Proof. If $\bar{x}$ happens to be a local minimizer of f then, as stated in Remark 3.4(2), one has $\mathbf{0}^* \in \mathcal{D}^{\triangle} f(\bar{x})$. In such event inequality (3.3) is trivially satisfied. Otherwise, fix an arbitrary $\epsilon > 0$. Corresponding to it, by definition of strong slope, there exists $\delta_\epsilon > 0$ such that

$$\sup_{x \in \mathrm{B}(\bar{x}, \delta_\epsilon) \setminus \{\bar{x}\}} \frac{f(\bar{x}) - f(x)}{d(x, \bar{x})} < |\nabla f|(\bar{x}) + \epsilon.$$

This means that $\bar{x}$ is a local minimizer of function $f(\cdot) + (|\nabla f|(\bar{x}) + \epsilon)\|\cdot - \bar{x}\|$. By applying again what has been noted in Remark 3.4(2), (3) and (4), one obtains

$$\begin{aligned} \mathbf{0}^* \quad &\in \quad \mathcal{D}^{\triangle}[f(\cdot) + (|\nabla f|(\bar{x}) + \epsilon)\|\cdot - \bar{x}\|](\bar{x}) \subseteq \mathcal{D}^{\triangle} f(\bar{x}) + (|\nabla f|(\bar{x}) + \epsilon)\mathcal{D}^{\triangle}\|\cdot - \bar{x}\|(\bar{x}) \\ &= \quad \mathcal{D}^{\triangle} f(\bar{x}) + (|\nabla f|(\bar{x}) + \epsilon)\mathbb{B}^*. \end{aligned}$$

Such an inclusion implies the existence of $v^* \in \mathcal{D}^{\triangle} f(\bar{x})$ such that $\|v^*\| \leq |\nabla f|(\bar{x}) + \epsilon$. This allows one to deduce that

$$\mathrm{dist}\left(\mathbf{0}^*, \mathcal{D}^{\triangle} f(\bar{x})\right) \leq |\nabla f|(\bar{x}) + \epsilon,$$

whence the thesis follows by arbitrariness of ϵ. □

3.2. **Lower exhausters of generalized derivatives.** Let $h : \mathbb{X} \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ be a positively homogeneous of degree one (henceforth, for short, p.h.) function. The need of extending the Minkowski-Hörmander duality in such a way to dually represent classes of p.h. functions, which are broader than $\mathcal{S}(\mathbb{X})$, led to introduce the notion of lower and upper families of exhausters (see [12]). Accordingly, a family $\underline{\mathcal{E}}(h) \subseteq \mathcal{K}(\mathbb{X}^*)$ is said to be a *lower exhauster* of h if

$$h(x) = \inf_{E \in \underline{\mathcal{E}}(h)} \max_{x^* \in E} \langle x^*, x \rangle = \inf_{E \in \underline{\mathcal{E}}(h)} \varsigma(x, E), \quad \forall x \in \mathbb{X}.$$

In nonsmooth optimization the p.h. functions to be dually represented are often generalized derivatives. Among them, in the investigations here exposed, the Hadamard directional derivatives will be employed. Given a function $f : \mathbb{X} \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ and $\bar{x} \in \mathrm{dom}\,(f)$, denote by

$$f^{\downarrow}_{\mathcal{H}}(\bar{x}; v) = \liminf_{\substack{v' \to v \\ t \to 0^+}} \frac{f(\bar{x} + tv') - f(\bar{x})}{t} \quad \text{and} \quad f^{\uparrow}_{\mathcal{H}}(\bar{x}; v) = \limsup_{\substack{v' \to v \\ t \to 0^+}} \frac{f(\bar{x} + tv') - f(\bar{x})}{t}$$

respectively the *Hadamard lower derivative* and the *Hadamard upper derivative* of f at $\bar{x}$, in the direction $v \in \mathbb{X}$. Notice that, if f is directionally differentiable at $\bar{x}$ in all directions and it is locally Lipschitz around the same point, then for every $v \in \mathbb{X}$ it is $f^{\downarrow}_{\mathcal{H}}(\bar{x}; v) = f^{\uparrow}_{\mathcal{H}}(\bar{x}; v) = f'(\bar{x}; v)$. In the subsequent section, when dealing with a function f, whose Hadamard lower derivative at $\bar{x}$ admits lower exhauster, the shortened notation

$$\mathcal{E}^{\downarrow}_{\mathcal{H}}(f, \bar{x}) = \underline{\mathcal{E}}(f^{\downarrow}_{\mathcal{H}}(\bar{x}; \cdot)).$$

will be used. In order to formulate a nondegeneracy condition involving lower exhausters of Hadamard lower derivatives, it is useful to define the 'norm' of a family $\underline{\mathcal{E}}(h)$ of elements in $\mathcal{K}(\mathbb{X}^*)$ as

$$\|\underline{\mathcal{E}}(h)\| = \sup_{E \in \underline{\mathcal{E}}(h)} \mathrm{dist}\,(\mathbf{0}^*, E)\,.$$

4. Sufficient conditions in Banach spaces

4.1. **Weak sharp minimality conditions for q.d. extremum problems.**

Theorem 4.1. *Let $(\mathbb{X}, \|\cdot\|)$ be a Banach space and let $f : \mathbb{X} \longrightarrow \mathbb{R}$ be a function l.s.c. on $\mathbb{X}$. Suppose that f is q.d. at each point of $\mathbb{X}$ and*

$$\mathrm{dist}\left(\mathbf{0}^*, \mathcal{D}^{\triangle} f(\mathbb{X} \setminus \mathrm{Argmin}(f))\right) = \tau > 0. \tag{4.1}$$

Then, f admits global weak sharp minimizers and $\mathrm{wsha}(f) \geq \tau$.

Proof. Since f is l.s.c. on $\mathbb{X}$, set $\mathbb{X}\backslash\mathrm{Argmin}(f)$ is open. By exploiting the estimate established in Lemma 3.6, one obtains in force of condition (4.1)

$$\inf_{x\in\mathbb{X}\backslash\mathrm{Argmin}(f)} |\nabla f|(x) \geq \inf_{x\in\mathbb{X}\backslash\mathrm{Argmin}(f)} \mathrm{dist}\left(\mathbf{0}^*, \mathcal{D}^{\triangle} f(x)\right) = \tau > 0.$$

Thus the nondegeneracy condition (2.6) of Theorem 2.4 appears to be fulfilled. This allows one to achieve both the assertions in the thesis. □

Theorem 4.1 extends to the broader class of q.d. functions a well-known sufficient condition for the global weak sharp minimality of convex functions (see, for instance, [31]), which is stated below.

Corollary 4.2. *Let $(\mathbb{X}, \|\cdot\|)$ be a Banach space and let $f : \mathbb{X} \longrightarrow \mathbb{R}\cup\{+\infty\}$ be a function l.s.c. and convex on $\mathbb{X}$. If*

$$\mathrm{dist}\,(\mathbf{0}^*, \partial f(\mathbb{X}\backslash\mathrm{Argmin}(f))) = \tau > 0, \tag{4.2}$$

then f admits global weak sharp minimizers and $\mathrm{wsha}(f) \geq \tau$.

Proof. If f is continuous on $\mathbb{X}$ it is enough to observe that, as remarked in Example 3.5, under the current assumptions it is $\mathcal{D}^{\triangle} f = \partial f$, and then to apply Theorem 4.1.

More in general, the Brøndsted-Rockafellar theorem is known to ensure the nonemptiness of the set valued map ∂f norm densely in $\mathrm{dom}\,(f)$. In the light of Example 2.3, at each subdifferentiable point it is $|\nabla f|(x) = \mathrm{dist}\,(\mathbf{0}^*, \partial f(x))$. In the case $x \notin \mathrm{dom}\,(\partial f)$, by convention it is $\mathrm{dist}\,(\mathbf{0}^*, \partial f(x)) = +\infty$. Thus, it is possible to apply directly Theorem 2.4, condition (2.6) being fufilled. □

Theorem 4.1 is not the only extension of the sufficient condition valid for the convex case. In fact, in [24] a similar result was presented, which relies on a nondegeneracy condition involving Fréchet subdifferentials. In contrast to Theorem 4.1, such result requires an Asplundity assumption on the underlying Banach space. On the other hand, in the former condition a generalized differentiability assumptions is made on f, while this can be avoided in [24], because nonemptiness of Fréchet subdifferentials is guaranteed norm densely by virtue of the Asplund property (see [22]). Thus these two conditions can not be obtained from the other each. Notice that the key step in the proof presented in [24] is the use of a fuzzy sum rule for Fréchet subdifferential valid in force of the Fréchet trustworthiness of the underlying space, the latter property being an equivalent manifestation of the Asplund one. All of this is not needed in the approach here exposed, since $\mathcal{D}^{\triangle}$ inherits from Clarke subdifferential calculus the sum rule adequate to the present circumstance (remember Remark 3.4(3)).

Since in context of q.d. functions condition (4.1) implies the nondegeneracy condition (2.6), the former can not be expected to be a characterization of global weak sharp minimality (remember Example 2.5). Nevertheless, an interesting consequence of condition (4.1), which is worth noting here, is that, in the presence of an additional semicontinuity assumption on the map $\mathcal{D}^{\triangle} f : \mathbb{X} \longrightarrow \mathcal{K}(\mathbb{X}^*)$, it prevents f to be smooth at any point of $\mathrm{bd}\,\mathrm{Argmin}(f)$, that is the boundary of $\mathrm{Argmin}(f)$, as proved below.

Proposition 4.3. *Let $(\mathbb{X}, \|\cdot\|)$ be a Banach space and let $f : \mathbb{X} \longrightarrow \mathbb{R}$ be a l.s.c. function, which is q.d. at each point of $\mathbb{X}$. Suppose that*

(i) condition (4.1) holds for some $\tau > 0$;

(ii) map $\mathcal{D}^{\triangle} f : \mathbb{X} \longrightarrow \mathcal{K}(\mathbb{X}^*)$ is norm-to-norm u.s.c. at each point of $\mathrm{Argmin}(f)$.

Then, f fails to be Gâteaux differentiable at each point of $\mathrm{bd}\,\mathrm{Argmin}(f)$.

Proof. According to Theorem 4.1, it is $\mathrm{Argmin}(f) \neq \varnothing$. If $\mathrm{Argmin}(f) \neq \mathbb{X}$ (that is f is not constant), since $\mathrm{Argmin}(f)$ is closed there exists $\bar{x} \in \mathrm{bd}\,\mathrm{Argmin}(f) \subseteq \mathrm{Argmin}(f)$. Assume ab absurdo that f is Gâteaux differentiable at $\bar{x}$. As remarked in Example 3.5, in such event it is $\mathcal{D}^{\triangle} f(\bar{x}) = \{\nabla f(\bar{x})\}$. As it is $\bar{x} \in \mathrm{Argmin}(f)$, the well-known Fermat rule implies $\mathcal{D}^{\triangle} f(\bar{x}) = \{\mathbf{0}^*\}$. By hypothesis (ii), corresponding to $\tau/2$ there exists $\delta_\tau > 0$ such that

$$\mathcal{D}^{\triangle} f(\mathrm{B}\,(\bar{x}, \delta_\tau)) \subseteq \mathrm{B}\left(\mathcal{D}^{\triangle} f(\bar{x}), \frac{\tau}{2}\right) = \frac{\tau}{2}\mathbb{B}^*,$$

where clearly $\mathrm{B}\,(A, r) = \mathrm{lev}_{\leq r}\,\mathrm{dist}\,(\cdot, A)$. Since $\mathrm{B}\,(\bar{x}, \delta_\tau) \cap (\mathbb{X}\backslash\mathrm{Argmin}(f)) \neq \varnothing$, the last inclusion contradicts condition (4.1), which is in force by (i). This completes the proof. □

Even though, as remarked above, condition (4.1) is only sufficient for global weak sharp minimality, Proposition 4.3 shows that for a certain class of functions global weak sharp minimality is incompatible with differentiability. To quote a metaphor due to V.F. Demyanov (see [13]), on the account of the nice properties enjoyed by weak sharp minimality, it is possible to assert that 'ugly ducklings' appear in this circumstance to play the role of 'beautiful swans'.

Let us pass now to the study of conditions for global weak sharp minimality in the context of constrained extremum problems. Given functions $f:\mathbb{X}\longrightarrow\mathbb{R}$, $g:\mathbb{X}\longrightarrow\mathbb{R}$ and $h:\mathbb{X}\longrightarrow\mathbb{R}$, here optimization problems of the form

$$(\mathcal{P}_c) \qquad \min_{x\in\mathbb{X}} f(x) \text{ subject to } g(x)\le 0,\ h(x)=0$$

will be considered. For convenience-sake the feasible region of $(\mathcal{P}_c)$ is denoted by Ω, i.e. $\Omega=\mathrm{lev}_{\le 0}\, g\cap h^{-1}(0)$, whereas the set of all its global solutions is indicated by $\mathrm{Argmin}(f,\Omega)$. If all data of $(\mathcal{P}_c)$ are functions, which are q.d. at each point of $\mathbb{X}$, while f and g are l.s.c., and h continuous on $\mathbb{X}$, then problem $(\mathcal{P}_c)$ will be referred to as a *q.d. problem*. Note that, whenever $(\mathcal{P}_c)$ is a q.d. problem, its feasible region is an example of what is called a q.d. set, according to [16].

In the context of constrained extremum problems the notion of global weak sharp minimality is modified as follows: a problem $(\mathcal{P}_c)$ is said to admit *constrained global weak sharp minimizers* if $\mathrm{Argmin}(f,\Omega)\neq\varnothing$ and there exists $\sigma>0$ such that

$$\sigma\cdot\mathrm{dist}\,(x,\mathrm{Argmin}(f,\Omega))\le f(x)-\inf_{x\in\Omega}f(x),\quad \forall x\in\Omega.$$

The supremum over all values σ satisfying the above inequality will be called *modulus of constrained global weak sharpness* of $(\mathcal{P}_c)$ and denoted by $\mathrm{wsha}(f,\Omega)$.

The next result provide a sufficient condition for global constrained minimizers to be weak sharp.

Theorem 4.4. *With reference to a q.d. problem $(\mathcal{P}_c)$, suppose that*

(i) $\mathrm{Argmin}(f,\Omega)\neq\varnothing$;

(ii) *there exists $\tau>0$ such that*

$$\inf_{x\in\mathbb{X}\backslash\Omega}\ \mathrm{dist}\left(\mathbf{0}^*,\mathcal{D}^{\triangle}[g]_+(x)+\mathcal{D}^{\triangle}|h|(x)\right)\ge\tau; \tag{4.3}$$

(iii) *function f is Lipschitz continuous on $\mathbb{X}$ with rank ℓ_f;*

(iv) *for some $\lambda>\ell_f$ and $\zeta>0$ it holds*

$$\inf_{x\in\mathbb{X}\backslash\mathrm{Argmin}(f,\Omega)}\mathrm{dist}\left(\mathbf{0}^*,\mathcal{D}^{\triangle}f(x)+\lambda\tau^{-1}\left(\mathcal{D}^{\triangle}[g]_+(x)+\mathcal{D}^{\triangle}|h|(x)\right)\right)\ge\zeta. \tag{4.4}$$

Then, the solutions of $(\mathcal{P}_c)$ are constrained global weak sharp minimizers and $\mathrm{wsha}(f,\Omega)\ge\zeta$.

Proof. First of all, observe that under the hypotheses made, since it is $\Omega\neq\varnothing$, then by virtue of Corollary 2.7 the following error bound holds true

$$\mathrm{dist}\,(x,\Omega)\le\tau^{-1}\left([g(x)]_++|h(x)|\right),\quad\forall x\in\mathbb{X}. \tag{4.5}$$

To prove this, notice that, being $(\mathcal{P}_c)$ q.d., function $[g]_++|h|$ is q.d. as well. Then, it suffices to recall that, according to Lemma 3.6, it results in

$$|\nabla([g]_++|h|)|(x)\ge\mathrm{dist}\left(\mathbf{0}^*,\mathcal{D}^{\triangle}([g]_++|h|)(x)\right)\ge\mathrm{dist}\left(\mathbf{0}^*,\mathcal{D}^{\triangle}[g]_+(x)+\mathcal{D}^{\triangle}|h|(x)\right),$$

as it is

$$\mathcal{D}^{\triangle}([g]_++|h|)(x)\subseteq\mathcal{D}^{\triangle}[g]_+(x)+\mathcal{D}^{\triangle}|h|(x)$$

and inequality (4.3) is in force. In the light of the validity of error bound (4.5), since f is Lipschitz continuous by hypothesis (iii), set Ω is closed and problem $(\mathcal{P}_c)$ admits global solutions, it is possible to invoke the basic principle of exact penalization [1]. According to it, if taking any $\ell>\ell_f$, it turns out that

$$\mathrm{Argmin}(f,\Omega)=\mathrm{Argmin}(f+\ell\tau^{-1}([g]_++|h|))\quad\text{and}\quad\inf_{x\in\Omega}f(x)=\inf_{x\in\mathbb{X}}[f(x)+\ell\tau^{-1}([g]_+(x)+|h|(x))],$$

[1] For a statement of such principle the reader is referred to [17] (see Theorem 6.8.1). It is to be noted that, in the mentioned reference, being formulated in finite dimensional spaces, the principle is proved by using the proximinality property enjoyed by any nonempty closed subset of $\mathbb{R}^n$. Nevertheless, with a slight modification, the proof can be rendered valid in any metric space.

where it is to be remarked that the right-side solution set in the first of the above equalities relates to an unconstrained optimization problem. The objective function of the latter is clearly l.s.c. and q.d. on $\mathbb{X}$ as well as bounded from below (remember that f is so). Consequently, one can apply Theorem 4.1, after having shown that the nondegeneracy condition (4.1) is fulfilled. This happens provided that $\ell = \lambda$ as in (4.4), because

$$\mathcal{D}\triangleq\left(f+\lambda\tau^{-1}([g]_+|h|\right)(x)\subseteq\mathcal{D}\triangleq f(x)+\lambda\tau^{-1}\left(\mathcal{D}\triangleq[g]_+(x)+\mathcal{D}\triangleq|h|(x)\right).$$

Thus, it results in

$$\zeta\cdot\mathrm{dist}\,(x,\mathrm{Argmin}(f,\Omega))\leq f(x)+\lambda\tau^{-1}\left([g]_+(x)+|h|(x)\right)-\inf_{x\in\Omega}f(x),\quad\forall x\in\mathbb{X},$$

which, for any $x\in\Omega$, gives the inequality to be proved. □

To complement the formulation of Theorem 4.4, conditions (4.3) and (4.4) should be expressed in terms of problem data, namely in terms of dual constructions directly related to functions g and h, not to $[g]_+$ and $|h|$, whose involvement is only instrumental. This can be done by virtue of the rich apparatus calculi, which is at disposal for q.d. functions. Notice that, in both cases, the question consists in computing $\mathcal{D}\triangleq[g]_+(x)+\mathcal{D}\triangleq|h|(x)$. This is carried out in the below Remark, under an additional continuity assumption on g, aimed at simplifying already involved formulæ.

Remark 4.5. Given $g:\mathbb{X}\longrightarrow\mathbb{R}$ and $h:\mathbb{X}\longrightarrow\mathbb{R}$, suppose that both are continuous and q.d. on $\mathbb{X}$, with

$$\mathcal{D}g(x)=[\underline{\partial}g(x),-\overline{\partial}g(x)]_\sim\quad\text{ and }\quad\mathcal{D}h(x)=[\underline{\partial}h(x),-\overline{\partial}h(x)]_\sim.$$

As one expects, the basic tool for calculating $\mathcal{D}\triangleq[g]_+(x)+\mathcal{D}\triangleq|h|(x)$ is the following well-known rule expressing the quasidifferential of a pointwise max-type function (see, for instance, [15, 16]). If functions $\phi:\mathbb{X}\longrightarrow\mathbb{R}$ and $\psi:\mathbb{X}\longrightarrow\mathbb{R}$ are q.d. at x, with

$$\mathcal{D}\phi(x)=[\underline{\partial}\phi(x),-\overline{\partial}\phi(x)]_\sim\quad\text{ and }\quad\mathcal{D}\psi(x)=[\underline{\partial}\psi(x),-\overline{\partial}\psi(x)]_\sim,$$

and if $\phi(x)=\psi(x)$, then, defined function $\phi\vee\psi:\mathbb{X}\longrightarrow\mathbb{R}$ as

$$(\psi\vee\psi)(x)=\max\{\phi(x),\,\psi(x)\},\quad x\in\mathbb{X},$$

the quasidifferential $\mathcal{D}(\phi\vee\psi)(x)=[\underline{\partial}(\phi\vee\psi)(x),-\overline{\partial}(\phi\vee\psi)(x)]_\sim$ is given according to the formulæ

$$\underline{\partial}(\phi\vee\psi)(x)=\mathrm{cl\,conv}\left[\left(\underline{\partial}\phi(x)-\overline{\partial}\psi(x)\right)\cup\left(\underline{\partial}\psi(x)-\overline{\partial}\phi(x)\right)\right]\quad\text{ and }\quad\overline{\partial}(\phi\vee\psi)(x)=\overline{\partial}\phi(x)+\overline{\partial}\psi(x).$$

Consequently, since it is $[g]_+=g\vee\mathbf{0}^*$, then if $x\in\mathbb{X}$ is such that $g(x)=0$, one obtains

$$\underline{\partial}[g]_+(x)=\mathrm{cl\,conv}\left[\underline{\partial}g(x)\cup(-\overline{\partial}g(x))\right]\quad\text{ and }\quad\overline{\partial}[g]_+(x)=\overline{\partial}g(x).\tag{4.6}$$

Analogously, since it is $|h|=h\vee(-h)$, then if $x\in\mathbb{X}$ is such that $h(x)=0$, one obtains

$$\underline{\partial}|h|(x)=2\,\mathrm{cl\,conv}\left[\underline{\partial}h(x)\cup(-\overline{\partial}h(x))\right]\quad\text{ and }\quad\overline{\partial}|h|(x)=\overline{\partial}h(x)-\underline{\partial}h(x).\tag{4.7}$$

Now, if $x\in\mathbb{X}\backslash\Omega$, the violation of the equality/inequality system defining Ω can be reduced to one of the following seven cases, which in turn leads to a different expression of set $\mathcal{D}\triangleq[g]_+(x)+\mathcal{D}\triangleq|h|(x)$, on the base of formulæ (4.6) and (4.7). Take into account that, because of the continuity of g and h, the nonnull sign of their values persists in a whole neighbourhood of a point at which such sign is taken.

case 1: $g(x)<0$ and $h(x)<0$. Since in this case locally it is $[g]_+=\mathbf{0}^*$ and $|h|=-h$, it results in

$$\mathcal{D}\triangleq[g]_+(x)+\mathcal{D}\triangleq|h|(x)=\left(-\overline{\partial}h(x)\right)\ \underline{\triangle}\ \underline{\partial}h(x);$$

case 2: $g(x)<0$ and $h(x)>0$. Since in this case locally it is $[g]_+=\mathbf{0}^*$ and $|h|=h$, it results in

$$\mathcal{D}\triangleq[g]_+(x)+\mathcal{D}\triangleq|h|(x)=\underline{\partial}h(x)\ \underline{\triangle}\ \left(-\overline{\partial}h(x)\right);$$

case 3: $g(x)=0$ and $h(x)<0$. Being in this case locally $|h|=-h$, by recalling formulæ (4.6) one finds

$$\mathcal{D}\triangleq[g]_+(x)+\mathcal{D}\triangleq|h|(x)=\left\{\mathrm{cl\,conv}\left[\underline{\partial}g(x)\cup(-\overline{\partial}g(x))\right]\ \underline{\triangle}\ \left(-\overline{\partial}g(x)\right)\right\}+\left[\left(-\overline{\partial}h(x)\right)\ \underline{\triangle}\ \underline{\partial}h(x)\right];$$

case 4: $g(x)=0$ and $h(x)>0$. Being in this case locally $|h|=h$, by recalling formulæ (4.6) one finds

$$\mathcal{D}\triangleq[g]_+(x)+\mathcal{D}\triangleq|h|(x)=\left\{\mathrm{cl\,conv}\left[\underline{\partial}g(x)\cup(-\overline{\partial}g(x))\right]\ \underline{\triangle}\ \left(-\overline{\partial}g(x)\right)\right\}+\left[\underline{\partial}h(x)\ \underline{\triangle}\ \left(-\overline{\partial}h(x)\right)\right];$$

case 5: $g(x) > 0$ and $h(x) < 0$. Being in this case locally $[g]_+ = g$ and $|h| = -h$, one obtains

$$\mathcal{D}^{\triangle}[g]_+(x) + \mathcal{D}^{\triangle}|h|(x) = \left[\underline{\partial}g(x) \,\underline{\triangle}\, \left(-\overline{\partial}g(x)\right)\right] + \left[\left(-\overline{\partial}h(x)\right) \,\underline{\triangle}\, \underline{\partial}h(x)\right];$$

case 6: $g(x) > 0$ and $h(x) > 0$. Being in this case locally $[g]_+ = g$ and $|h| = h$, one obtains

$$\mathcal{D}^{\triangle}[g]_+(x) + \mathcal{D}^{\triangle}|h|(x) = \left[\underline{\partial}g(x) \,\underline{\triangle}\, \left(-\overline{\partial}g(x)\right)\right] + \left[\underline{\partial}h(x) \,\underline{\triangle}\, \left(-\overline{\partial}h(x)\right)\right];$$

case 7: $g(x) > 0$ and $h(x) = 0$. In this case, locally it is $[g]_+ = g$. Thus, by recalling formulæ (4.7), one gets

$$\mathcal{D}^{\triangle}[g]_+(x) + \mathcal{D}^{\triangle}|h|(x) = \left[\underline{\partial}g(x) \,\underline{\triangle}\, \left(-\overline{\partial}g(x)\right)\right] + \left\{2\,\mathrm{cl\,conv}\,\left[\underline{\partial}h(x) \cup (-\overline{\partial}h(x))\right] \,\underline{\triangle}\, \left(\underline{\partial}h(x) - \overline{\partial}h(x)\right)\right\}.$$

Notice that, in order to calculate $\mathcal{D}^{\triangle}[g]_+(x) + \mathcal{D}^{\triangle}|h|(x)$ as it appears in condition (4.4), a further distinction is needed, according to the fact that $x \in \Omega$ or $x \notin \Omega$. Since this can be carried out as done above, the details are omitted.

4.2. **A weak sharp minimality condition in terms of exhausters.**

Theorem 4.6. *Let $(\mathbb{X}, \|\cdot\|)$ be a Banach space and let $f : \mathbb{X} \longrightarrow \mathbb{R}$ be a function l.s.c. on $\mathbb{X}$. Suppose that $f^{\downarrow}_{\mathcal{H}}(x;\cdot)$ admits a lower exhauster $\mathcal{E}^{\downarrow}_{\mathcal{H}}(f,x)$ at each point of $\mathbb{X}$ and*

$$\inf_{x \in \mathbb{X}\setminus \mathrm{Argmin}(f)} \|\mathcal{E}^{\downarrow}_{\mathcal{H}}(f,x)\| = \tau > 0. \tag{4.8}$$

Then, f admits global weak sharp minimizers and $\mathrm{wsha}(f) \geq \tau$.

Proof. Ab absurdo, assume that f has no global weak sharp minimizers. This amounts to suppose that either $\mathrm{Argmin}(f) = \varnothing$ or inequality (1.1) is not valid. In both cases, according to the characterization provided by Theorem 2.1 there must exist $\hat{x} \in \mathbb{X}$ and $\hat{\alpha} > \underline{\alpha}$ such that $\hat{x} \notin \mathrm{lev}_{\leq \hat{\alpha}} f$ and

$$f(\hat{x}) < \underline{\alpha} + \tau \cdot \mathrm{dist}\,(\hat{x}, \mathrm{lev}_{\leq \hat{\alpha}} f).$$

Since this inequality is strict, it is possible to find $\epsilon > 0$ such that

$$\epsilon < \min\{\tau, \mathrm{dist}\,(\hat{x}, \mathrm{lev}_{\leq \hat{\alpha}} f)\}$$

and

$$f(\hat{x}) < \underline{\alpha} + (\tau - \epsilon)(\mathrm{dist}\,(\hat{x}, \mathrm{lev}_{\leq \hat{\alpha}} f) - \epsilon).$$

Thus, the Ekeland variational principle ensures the existence of $\bar{x} \in \mathbb{X}$ such that

$$d(\bar{x}, \hat{x}) \leq \mathrm{dist}\,(\hat{x}, \mathrm{lev}_{\leq \hat{\alpha}} f) - \epsilon \tag{4.9}$$

and

$$f(\bar{x}) < f(x) + (\tau - \epsilon)\|x - \bar{x}\|, \quad \forall x \in \mathbb{X}\setminus\{\bar{x}\}.$$

As a consequence of the fact that $\bar{x}$ (globally) minimizes function $f(\cdot) + (\tau - \epsilon)\|\cdot - \bar{x}\|$, one obtains

$$\begin{aligned} 0 &\leq (f(\cdot) + (\tau - \epsilon)\|\cdot - \bar{x}\|)^{\downarrow}_{\mathcal{H}}(\bar{x}; v) \leq f^{\downarrow}_{\mathcal{H}}(\bar{x}; v) + (\tau - \epsilon)\|\cdot - \bar{x}\|^{\uparrow}_{\mathcal{H}}(\bar{x}; v) \\ &= f^{\downarrow}_{\mathcal{H}}(\bar{x}; v) + (\tau - \epsilon)\|\cdot - \bar{x}\|'(\bar{x}; v), \quad \forall v \in \mathbb{X} \end{aligned}$$

(remember that $\|\cdot - \bar{x}\|$ is Lipschitz continuous and directionally differentiable at $\bar{x}$). By hypothesis $f^{\downarrow}_{\mathcal{H}}(\bar{x};\cdot)$ admits a lower exhauster $\mathcal{E}^{\downarrow}_{\mathcal{H}}(f,\bar{x})$. Thus, being

$$\mathcal{E}^{\downarrow}_{\mathcal{H}}((\tau - \epsilon)\|\cdot - \bar{x}\|, \bar{x}) = \{(\tau - \epsilon)\mathbb{B}^*\},$$

from the last inequalities it follows

$$0 \leq \inf_{E \in \mathcal{E}^{\downarrow}_{\mathcal{H}}(f,\bar{x})} \varsigma(v, E) + (\tau - \epsilon)\varsigma(v, \mathbb{B}^*), \quad \forall v \in \mathbb{X}.$$

The last inequality implies in the light of the Minkowski-Hörmander duality that for every $E \in \mathcal{E}^{\downarrow}_{\mathcal{H}}(f,\bar{x})$ it holds

$$0 \leq \varsigma(v, E) + (\tau - \epsilon)\varsigma(v, \mathbb{B}^*) = \varsigma(v, E + (\tau - \epsilon)\mathbb{B}^*), \quad \forall v \in \mathbb{X}.$$

Consequently, one finds

$$\mathbf{0}^* \in E + (\tau - \epsilon)\mathbb{B}^*, \quad \forall E \in \mathcal{E}^{\downarrow}_{\mathcal{H}}(f,\bar{x}).$$

This inclusion implies that in every $E \in \mathcal{E}^{\downarrow}_{\mathcal{H}}(f, \bar{x})$ there must exist an element v^*_E such that $\|v^*_E\| \le \tau - \epsilon$, so that $\mathrm{dist}\,(\mathbf{0}^*, E) \le \tau - \epsilon$. Thus, it is

$$\sup_{E \in \mathcal{E}^{\downarrow}_{\mathcal{H}}(f,\bar{x})} \mathrm{dist}\,(\mathbf{0}^*, E) \le \tau - \epsilon,$$

what contradicts hypothesis (4.8), if $x \notin \mathrm{Argmin}(f)$. Therefore one is forced to deduce that $\bar{x} \in \mathrm{Argmin}(f)(\bar{x}) \neq \varnothing$. To reach the concluding contradiction it is sufficient to observe that, by virtue of inequality (4.9), it is

$$d(\hat{x}, \mathrm{Argmin}(f)) \le \mathrm{dist}\,(\hat{x}, \mathrm{lev}_{\le \hat{\alpha}} f) - \epsilon,$$

whereas, being $\mathrm{Argmin}(f) \subseteq \mathrm{lev}_{\le \hat{\alpha}} f$, it should be at the same time

$$d(\hat{x}, \mathrm{Argmin}(f)) \ge \mathrm{dist}\,(\hat{x}, \mathrm{lev}_{\le \hat{\alpha}} f)\,.$$

Again the characterization provided by Theorem 2.1 allows to obtain the estimate of $\mathrm{wsha}(f)$ asserted in the thesis. □

Remark 4.7. Theorem 4.6 can be regarded as a further extension to nonconvex functions of Corollary 4.2. In fact, it applies to a class of functions that may not be q.d.. Concerning such class, it is worth noting that every function defined on a normed vector space $(\mathbb{X}, \|\cdot\|)$ and locally Lipschitz around a reference point $x \in \mathbb{X}$ has Hadamard derivatives at x admitting lower exhausters (see, for instance, [7]). Again, if a Banach space $(\mathbb{X}, \|\cdot\|)$ is locally uniformly convex, i.e.

$$\forall u \in \mathbb{S},\ \forall \epsilon > 0 \quad \text{it holds} \quad \sup_{v \in \mathbb{S} \setminus \mathrm{int}\,\mathrm{B}(u,\epsilon)} \left\| \frac{u+v}{2} \right\| < 1,$$

every function $f : \mathbb{X} \longrightarrow \mathbb{R}$, whose Hadamard lower derivative at x is continuous with respect to the argument direction, admits such class as $\mathcal{E}^{\downarrow}_{\mathcal{H}}(f, x)$, as a consequence of known results on the dual representation of p.h. functions (see, for instance, [29]). The class of all locally uniformly convex Banach spaces is actually rather wide, as it includes for example all reflexive Banach spaces.

The next result provides a condition for constrained global solutions of $(\mathcal{P}_c)$ to be weak sharp minimizers, in terms of lower exhausters. In this case, the constraints are left in a geometrical form, in the sense that they are not formalized by an equality/inequality system. In what follows, by $N^\circ(\Omega, x)$ the Clarke normal cone to set Ω at a point $x \in \mathbb{X}$ is denoted. In the case $x \notin \Omega$ set $N^\circ(\Omega, x) = \varnothing$.

Theorem 4.8. *Let $(\mathbb{X}, \|\cdot\|)$ be a Banach space, let $f : \mathbb{X} \longrightarrow \mathbb{R}$ be a Lipschitz continuous function on $\mathbb{X}$, with rank $\ell_f > 0$, and let $\Omega \subseteq \mathbb{X}$ be a nonempty closed set. Suppose that*

(i) $\mathrm{Argmin}(f, \Omega) \neq \emptyset$;

(ii) for some $\lambda > \ell_f$ it is

$$\inf_{x \in \Omega \setminus \mathrm{Argmin}(f,\Omega)} \ \sup_{E \in \mathcal{E}^{\downarrow}_{\mathcal{H}}(f,x)} \mathrm{dist}\,(\mathbf{0}^*, E + \lambda(N^\circ(\Omega, x) \cap \mathbb{B}^*)) = \zeta > 0. \tag{4.10}$$

Then, the solutions of $(\mathcal{P}_c)$ are constrained global weak sharp minimizers and $\mathrm{wsha}(f, \Omega) \ge \zeta$.

Proof. As in the constrained case already treated, let us start with observing that, under the current assumptions, one can reduce the constrained problem $(\mathcal{P}_c)$ to an unconstrained one. Indeed, since f is Lipschitz continuous, with rank ℓ_f, and $\mathrm{Argmin}(f, \Omega) \neq \emptyset$, then according to the basic penalization principle one has that, for every $\ell > \ell_f$, it holds

$$\mathrm{Argmin}(f, \Omega) = \mathrm{Argmin}(f + \ell \mathrm{dist}\,(\cdot, \Omega)).$$

Henceforth it is possible to follow the argument proposed in the proof of Theorem 4.6, with function f replaced here by $f + \lambda \mathrm{dist}\,(\cdot, \Omega)$. To this regard, observe that, since function f is Lipschitz continuous, its Hadamard lower derivative admits lower (also upper) exhausters at each point of $\mathbb{X}$. In the present case, since $\mathrm{Argmin}(f, \Omega) \neq \emptyset$, one needs only to assume ab absurdo that inequality (2.1) is not true. Thus, fixed $\hat{x} \in \mathbb{X}$ and $\epsilon > 0$, by proceeding as in the proof of Theorem 4.6 one gets the existence of $\bar{x} \in \mathbb{X}$ such that

$$d(\bar{x}, \hat{x}) \le \mathrm{dist}\,(\hat{x}, \mathrm{lev}_{\le \hat{\alpha}}\,(f + \lambda \mathrm{dist}\,(\cdot, \Omega))) - \epsilon$$

and

$$\begin{aligned} 0 &\le (f(\cdot)+\lambda\mathrm{dist}\,(\cdot,\Omega)+(\zeta-\epsilon)\|\cdot-\bar{x}\|)^{\downarrow}_{\mathcal{H}}(\bar{x};v) \\ &\le f^{\downarrow}_{\mathcal{H}}(\bar{x};v)+\mathrm{dist}\,(\cdot,\Omega)^{\uparrow}_{\mathcal{H}}(\bar{x};v)+(\zeta-\epsilon)\|\cdot-\bar{x}\|'(\bar{x};v) \\ &\le f^{\downarrow}_{\mathcal{H}}(\bar{x};v)+\mathrm{dist}\,(\cdot,\Omega)^{\circ}\,(\bar{x};v)+(\zeta-\epsilon)\|\cdot-\bar{x}\|'(\bar{x};v), \quad \forall v\in\mathbb{X}, \end{aligned}$$

where $\mathrm{dist}\,(\cdot,\Omega)^{\circ}\,(\bar{x};v)$ denotes the Clarke derivative of function $\mathrm{dist}\,(\cdot,\Omega)$ at $\bar{x}$ in the direction v. By recalling the dual representation

$$\mathrm{dist}\,(\cdot,\Omega)^{\circ}\,(\bar{x};v)=\max_{x^*\in\partial^\circ\mathrm{dist}(\cdot,\Omega)(\bar{x})}\langle x^*,v\rangle, \quad \forall v\in\mathbb{X},$$

and the fact that

$$\partial^\circ\mathrm{dist}\,(\cdot,\Omega)\,(\bar{x})=N^\circ(\Omega,\bar{x})\cap\mathbb{B}^*$$

(see, for instance, [8]), from the last inequalities one obtains

$$0\le\inf_{E\in\mathcal{E}^{\downarrow}_{\mathcal{H}}(f,\bar{x})}\varsigma(v,E)+\varsigma(v,\lambda(N^\circ(\Omega,\bar{x})\cap\mathbb{B}^*))+(\zeta-\epsilon)\varsigma(v,\mathbb{B}^*), \quad \forall v\in\mathbb{X}.$$

Thus, for every $E\in\mathcal{E}^{\downarrow}_{\mathcal{H}}(f,\bar{x})$ one has

$$\mathbf{0}^*\in E+\lambda(N^\circ(\Omega,\bar{x})\cap\mathbb{B}^*)+(\zeta-\epsilon)\mathbb{B}^*.$$

The last inclusion implies the fact that $\bar{x}\in\Omega$ and the existence of $v^*\in E+\lambda(N^\circ(\Omega,\bar{x})\cap\mathbb{B}^*)$, with $\|v^*\|\le\zeta-\epsilon$, what leads to deduce that $\bar{x}\in\mathrm{Argmin}(f,\Omega)$. This conclusion, if reasoning as in the proof of Theorem 4.6, allows one to reach the desired contradiction. □

References

1. Azé, D., Corvellec, J.-N., and Lucchetti, R.E., *Variational pairs and applications to stability in nonsmooth analysis*, Nonlinear Anal. **49** (2002), no. 5, 643–670.
2. Bednarczuk, E., *On weak sharp minima in vector optimization with applications to parametric problems*, Control Cybernet. **36** (2007), no. 3, 563–570.
3. Burke, J.V. and Deng, S., *Weak sharp minima revisited, part I: basic theory*, Well-posedness in optimization and related topics (Warsaw, 2001). Control Cybernet. **31** (2002), no. 3, 439–469.
4. Burke, J.V. and Deng, S., *Weak sharp minima revisited, part II: application to linear regularity and error bounds*, Math. Program. **104** (2005), no. 2-3, Ser. B, 235–261.
5. Burke, J.V. and Deng, S., *Weak sharp minima revisited. III. Error bounds for differentiable convex inclusions*, Math. Program. **116** (2009), no. 1-2, Ser. B, 37–56.
6. Burke, J.V. and Ferris, M.C., *Weak sharp minima in mathematical programming*, SIAM J. Control. Optim. **31** (1993), no. 5, 1340–1359.
7. Castellani, M., *A dual respresentation for proper positively homogeneous functions*, J. Global Optim. **16** (2000), no. 4, 393–400.
8. Clarke, F.H., *Optimization and Nonsmooth Analysis*, Wiley Interscience, New York, 1983.
9. Covitz, H. and Nadler, S.B.Jr., *Multi-valued contraction mappings in generalized metric spaces*, Israel J. Math. **8** (1970), 5–11.
10. De Giorgi, E., Marino, A., and Tosques, M., *Problems of evolution in metric spaces and maximal decreasing curves*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8) **68** (1980), no. 3, 180–187 [in Italian].
11. Demyanov, V.F., *Connections between a Clarke-subdifferential and a quasidifferential*, Vestnik Leningradskogo Universiteta **13** (1980), 18–24 [in Russian].
12. Demyanov, V.F., *Exhausters of a positively homogeneous function*, Optimization **45** (1999), no. 1-4, 13–29.
13. Demyanov, V.F., *Nonsmooth functions, who are they: Ugly ducklings or beautiful swans?*, Sankt-Peterburgskii Universitet **14** (2012), 17–19 [in Russian].
14. Demyanov, V. F. and Rubinov, A. M., *Quasidifferentiable functionals*, Dokl. Akad. Nauk SSSR **250** (1980), 21–25 [in Russian].
15. Demyanov, V.F. and Rubinov, A.M., *Quasidifferential Calculus*, Optmization Software, New York, 1986.
16. Demyanov, V.F. and Rubinov, A. M., *Constructive nonsmooth analysis*, Approximation & Optimization, 7, Peter Lang, Frankfurt am Main, 1995.
17. Facchinei, F. and Pang, J.-S., *Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume I*, Springer-Verlag, New York, 2003.
18. Gao, Y., *Demyanov difference of two sets and optimality conditions of Lagrange multiplier type for constrained quasidifferentiable optimization*, J. Optim. Theory Appl. **104** (2000), no. 2, 377–394.

19. Ioffe, A.D., *Necessary and sufficient conditions for a local minimum. I: a reduction theorem and first order conditions*, SIAM J. Control. Optim. **17** (1979), no. 2, 245–250.
20. Ioffe, A.D., *Metric regularity and subdifferential calculus*, Russian Math. Surveys **55** (2000), no. 3, 501–558.
21. Klatte, D. and Kummer, B., *Nonsmooth equations in optimization. Regularity, calculus, methods and applications*, Kluwer Academic Publishers, Dordrecht, 2002.
22. Mordukhovich, B.S., *Variational Analysis and Generalized Differentiation, I: Basic Theory*, Springer, Berlin, 2006.
23. Nadler, S.B.Jr., *Multi-valued contraction mappings*, Pacific J. Math. **30** (1969), 475–488.
24. Ng, K.F. and Zheng, X.Y., *Global weak sharp minima on Banach spaces*, SIAM J. Control Optim. **41** (2003), no. 6, 1868–1885.
25. Rubinov, A.M., *Differences of convex compact sets and their applications in nonsmooth analysis*, in Giannessi, F. (ed.) Nonsmooth Optimization. Methods and Applications, Gordon and Breach, Amsterdam (1992), 366–378.
26. Rubinov, A.M. and Akhundov, I.S., *Differences of compact sets in the sense of Demyanov and its applications to nonsmooth analysis*, Optimization **23** (1992), 179–188.
27. Rubinov, A.M. and Vladimirov, A.A., *Differences of convex compacta and metric spaces of convex compacta with applications: a survey*, in Demyanov, V.F., Rubinov, A.M. (eds.) Quasidifferentiability and related topics, Kluwer, Dordrecht (2000), 263–296.
28. Studniarski, M. and Ward, D.E., *Weak sharp minima: characterizations and sufficient conditions*, SIAM J. Control Optim. **38** (1999), no. 1, 219–236.
29. Uderzo, A., *Convex approximators, convexificators and exhausters: applications to constrained extremum problems*, in Demyanov, V.F., Rubinov, A.M. (eds.) Quasidifferentiability and related topics, Kluwer Academic Publishers, Dordrecht (2000), 297–327.
30. Uderzo A., *Convex Difference Criteria for the Quantitative Stability of Parametric Quasidifferentiable Systems*, Set-Valued Anal. **15** (2007), no. 1, 81–104.
31. Zălinescu, C., *Weak sharp minima, well-behaving functions and global error bounds for convex inequalities in Banach spaces*, in Bulatov, A. Baturin V. (eds.) Proceedings of the 12-th Baikal International Conference "Optimization methods and their applications", Institute of System Dynamics and Control Theory of SB RAS, Irkutsk (2001), 272–284.

A. Uderzo: Dept. of Mathematics and Applications, Università di Milano-Bicocca, Via Cozzi, 53 - 20125 Milano, Italy, e-mail: amos.uderzo@unimib.it